JVI Journals.ASM.org

Large Tailed Spindle Viruses of *Archaea*: a New Way of Doing Viral Business

Rebecca Hochstein,[a] Daniel Bollschweiler,[b] Harald Engelhardt,[b] C. Martin Lawrence,[b,c] Mark Young[d]

Departments of Microbiology and Immunology,[a] Chemistry and Biochemistry,[c] and Plant Sciences and Plant Pathology,[d] Montana State University, Bozeman, Montana, USA; Department of Molecular Structural Biology, Max-Planck-Institute for Biochemistry, Martinsried, Germany[b]

Viruses of *Archaea* continue to surprise us. Archaeal viruses have revealed new morphologies, protein folds, and gene content. This is especially true for large spindle viruses, which infect only *Archaea*. We present a comparison of particle morphologies, major coat protein structures, and gene content among the five characterized large spindle viruses to elucidate defining characteristics. Structural similarities and a core set of genes support the grouping of the large spindle viruses into a new superfamily.

Viruses infecting hosts in the third domain of life (*Archaea*) have long fascinated virologists due to their unusual morphologies and unique genetic content (1, 2). This is particularly true for the large tailed spindle viruses, which infect only *Archaea*. These viruses package circular double-stranded DNA (dsDNA) genomes within spindle (or lemon)-shaped virions and have tails protruding from one or both ends. To date, five of these large spindle viruses have been isolated from crenarchaeal hosts replicating in high-temperature hot springs around the world. These include the *Acidianus* two-tailed virus (ATV), *Sulfolobus tengchongensis* spindle-shaped virus 1 (STSV1) and STSV2, and *Sulfolobus* monocaudavirus 1 (SMV1) (3–7). The fifth member is a new virus, *Acidianus* tailed spindle virus (ATSV; proposed name), which we recently isolated from Crater Hills Alice Spring, a high-temperature (80°C) acidic (pH 2) hot spring in Yellowstone National Park (USA) (Fig. 1A). The 71-kb circular dsDNA genome of ATSV shares ~25% of its genes with STSV1 and -2 and a smaller subset with the other large spindle viruses. We and others (8) propose that large spindle viruses are distinguished from the smaller spindle-shaped *Fuselloviridae* archaeal viruses based on fundamental differences in their major structural proteins, genome content, and overall virion structures.

The virion structures of large spindle viruses are often pleomorphic, and for each specific virus, the dimensions of the central lemon-shaped capsid can vary in width and length. For example, ATSV dimensions range from 69 by 133 nm to 138 by 221 nm. The ATSV tails also vary greatly in length, ranging from 35 to 720 nm. Tail lengths are also variable in STSV1 and -2 and SMV1. In contrast to these single-tailed viruses, ATV is initially released from the infected cell as a tailless virion (3). Remarkably, extracellular ATV then develops two elongated tails extending from opposite ends of the spindle-shaped head. In rare cases, two-tailed particles have also been observed for some of the single-tailed viruses (4, 6). However, there is no evidence yet that these tails extend once progeny virions are released from the infected cell. Importantly, cryoelectron tomography suggests that unlike the classical head-tailed structures of some bacteriophage, the tails of these large spindle viruses are a continuous structure with the spindle head (Fig. 1A). Finally, both ATV and ATSV have 5-nm-width tail fibers extending 15 to 25 nm from the tips of their tails that are likely involved with attachment to their host cell surface. While tail fibers have yet to be reported for the others, STSV1 and -2 attach to host cells at the tip of their tails, indicating that there is some sort of attachment structure present (4, 5).

Despite some differences in virion structure, all five large spindle viruses share a major coat protein (MCP), with the ATSV MCP showing >39% identity to each of the other MCPs over the N-terminal half of the protein (Fig. 1B). We have recently determined the crystal structure of the ATSV MCP (Fig. 1C). The structure revealed a right-handed, antiparallel 4-helix bundle lacking overhand connections, with a fifth C-terminal helix coming off the helical bundle at a 50° angle. In addition, an extended loop connects the third and fourth helices. The fold of this structure is quite similar to that of the ATV MCP, P131 (9). In ATV, however, the fourth helix is kinked, and the fifth helix extends off the 4-helix bundle at a different angle than the fifth helix of ATSV. Sequence alignment of the MCPs of the 5 large spindle viruses indicate that helices 1 and 2 are well conserved but that the C terminus is more variable (Fig. 1B). However, there is close structural alignment of the ATSV and ATV proteins despite their lack of C-terminal amino acid identity (Fig. 1B and C). Thus, the MCPs of STSV1 and -2 and SMV1 are also expected to adopt a right-handed, antiparallel 4-helix bundle motif, similar to those of ATSV and ATV. Because it is the dominant protein in the virion, and because the ATSV virion appears as a continuous structure that smoothly transitions from head to tail, we propose that the MCP of ATSV contributes to the structure of both the spindle and the tail. How these helical proteins might form a spindle-shaped virus is an outstanding question in structural virology, with important implications for virus assembly, maturation, and infection.

Among *Archaea*, this MCP topology appears unique to the large spindle viruses. While the *Fuselloviridae* share similar spindle shapes, their MCPs (e.g., SSV1 VP1) are smaller and predicted to have two conserved transmembrane domains (8). Interestingly,

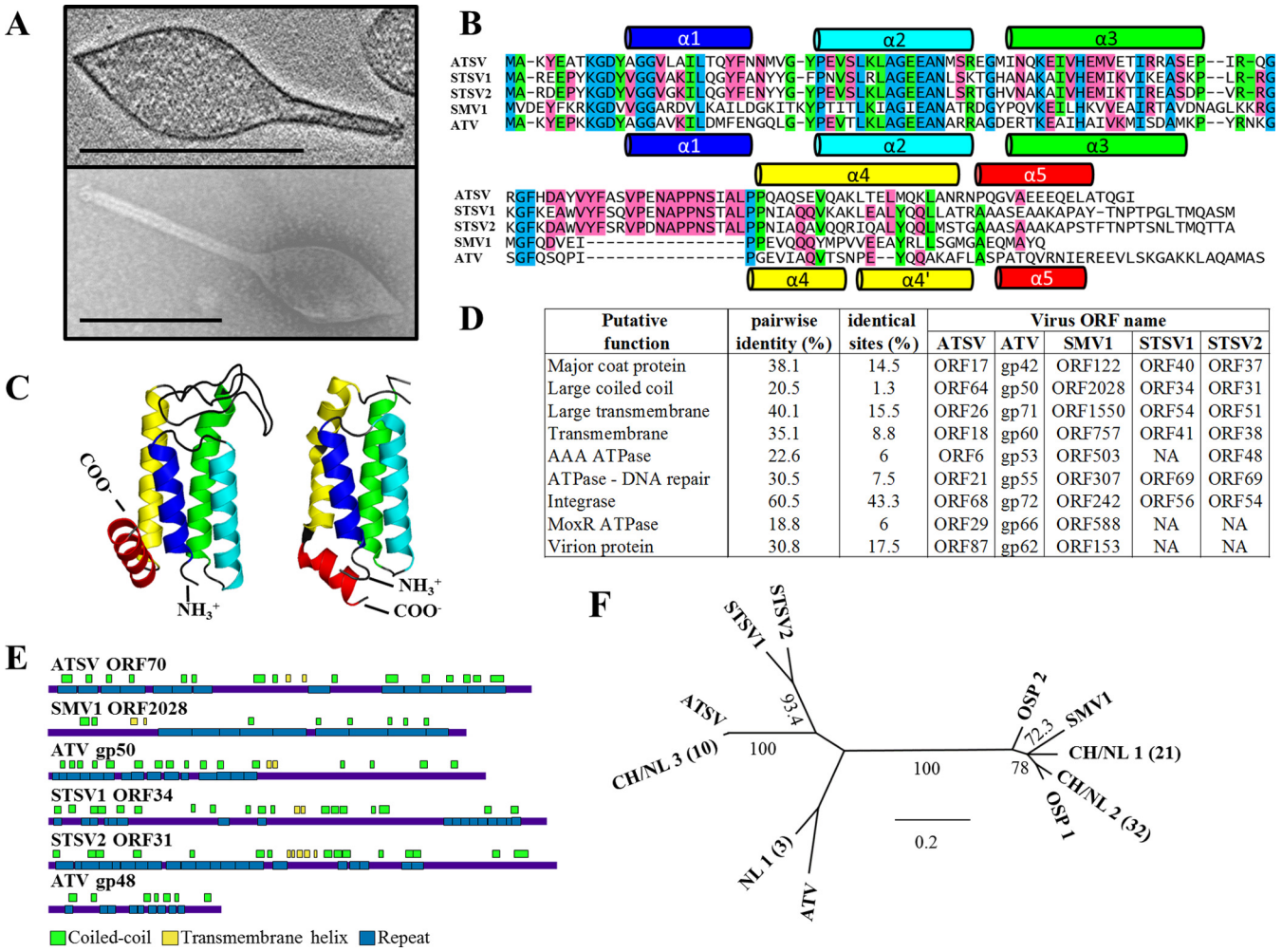
Accepted manuscript posted online 17 June 2015

Citation Hochstein R, Bollschweiler D, Engelhardt H, Lawrence CM, Young M. 2015. Large tailed spindle viruses of *Archaea*: a new way of doing viral business. J Virol 89:9146–9149. doi:10.1128/JVI.00612-15.

Editor: J. Pfeiffer

Address correspondence to Mark Young, myoung@montana.edu.

doi:10.1128/JVI.00612-15



Putative function	pairwise identity (%)	identical sites (%)	Virus ORF name				
			ATSV	ATV	SMV1	STSV1	STSV2
Major coat protein	38.1	14.5	ORF17	gp42	ORF122	ORF40	ORF37
Large coiled coil	20.5	1.3	ORF64	gp50	ORF2028	ORF34	ORF31
Large transmembrane	40.1	15.5	ORF26	gp71	ORF1550	ORF54	ORF51
Transmembrane	35.1	8.8	ORF18	gp60	ORF757	ORF41	ORF38
AAA ATPase	22.6	6	ORF6	gp53	ORF503	NA	ORF48
ATPase - DNA repair	30.5	7.5	ORF21	gp55	ORF307	ORF69	ORF69
Integrase	60.5	43.3	ORF68	gp72	ORF242	ORF56	ORF54
MoxR ATPase	18.8	6	ORF29	gp66	ORF588	NA	NA
Virion protein	30.8	17.5	ORF87	gp62	ORF153	NA	NA

FIG 1 ATSV virion and MCP structure and genomic analysis of large spindle viruses. (A) Central section through tomogram (top) of an unstained vitrified ATSV showing a continuous structure throughout the body and tail and negatively stained electron micrograph (bottom) showing variable tail length. Scale bars, 200 nm. (B) MUSCLE alignment of large spindle virus major coat proteins with ATSV and ATV helices from crystal structure. (C) Crystal structures of ATSV (left) and ATV (right). Blue, N terminus; red, C terminus. (D) Shared large spindle virus proteins. Average percent pairwise identity and percent identical sites were determined by MUSCLE alignment. (E) Large coiled-coil proteins with features highlighted, as predicted by Coils (coiled coils), RADAR (rapid automatic detection and alignment of repeats in protein sequences) (repeats), and TMHMM (transmembrane helices). (F) Neighbor-joining tree of MUSCLE alignment of large spindle virus MCPs and homologs from Yellowstone hot spring viral metagenomes showing a high level of diversity of large spindle virus types. Results are shown for Nymph Lake hot springs 10, 17, and 18 (NL), Crater Hills Alice hot spring (CH), and One Hundred Spring Plains hot spring (OSP). Number in parentheses indicates the number of sequences in group. The scale bar represents 0.2 amino acid substitutions per site. Bootstrap values from 1,000 replicates are shown.

structural work on *Sulfolobus islandicus* rod-shaped virus (10) and *Acidianus* filamentous virus (11) reveals MCPs with antiparallel 4-helix bundles at their core. However, these are left-handed 4-helix bundles rather than the right-handed bundles seen for the large-tailed spindle viruses (see reference 12 for topological definitions).

In addition to a shared structure for their MCPs, the large spindle viruses share similar genome structures and sizes. They have the largest known genomes among archaeal viruses, between 48 and 76 kbp, and contain 51 to 96 open reading frames (ORFs). All are circular, double-stranded DNA genomes. STSV1 has a single putative origin of replication based on a high AT content and tandem repeats (4), but similar replication origins for the other viruses have not been identified. Overall, little is known about the replication strategy of large spindle viruses.

Of the 51 to 96 ORFs in each genome, a core set of 9 ORFs is shared among all or most viruses. The core gene set encodes six proteins that are shared by all five viruses, one that is shared by all but STSV1, and two that are shared by ATSV, ATV, and SMV1 (Fig. 1D). The products of six of the core genes can be assigned putative functions. These include three ATPases (MoxR-like ATPase, AAA+ ATPase, and DnaA-like ATPase) (13) and an integrase from the tyrosine recombinase family. Two virus structural proteins, including the MCP, are also recognized. The three proteins lacking functional annotation have some interesting characteristics suggestive of function, including heptad motifs, uncharacterized repeats, and putative transmembrane domains.

The largest of these proteins (ATSV ORF64, ATV gp50, SMV1 ORF2028, STSV1 ORF34, and STSV2 ORF31) range in size from 212 to 257 kDa (1,940 to 2,235 amino acids), and secondary-struc-

ture predictions suggest a mostly helical structure. These proteins share multiple strings of heptad repeats, indicative of coiled-coil motifs (Fig. 1E), suggesting that they may form a homooligomeric quaternary structure. These proteins also contain longer, loosely conserved repeats with conserved charged and hydrophobic residues. Additionally, each protein contains at least two predicted transmembrane domains, usually in the middle of the protein (Fig. 1E). The large protein is one of 4 minor structural proteins detected in STSV1 virions and the only additional protein found in STSV2 virions besides the major structural protein. By analogy, it is likely that this protein is also a component of the other virions. With the large size and coiled-coil nature of these proteins, they are likely to be involved in the virus structure.

Cryoelectron tomography of ATV shows a 2-nm filament extending through the tails (3). The filament is thought to be composed of ATV gp48 protein, another coiled-coil- and repeat-containing virion protein found only in ATV (Fig. 1E). However, although ATV gp48 forms filaments and interacts with other proteins thought to be involved in tail formation, similar experiments have not been performed with the ATV large coiled-coil protein (ATV gp50). Given the similarities, the large coiled-coil proteins could provide a similar tail development function in the other large spindle viruses and maybe in ATV as well. Additionally, the tail fiber structure at the end of the tail in ATV and ATSV (Fig. 1A) could be composed of the ends of such a filament structure, similar to the unfurling of microtubules (3). Coiled-coil and other repeat motifs have been seen in virus attachment proteins such as T4 fibritin and hemagglutinin (14). It is reasonable that these large coiled-coil proteins could be involved in virus attachment to the host cell as part of the terminal tail fiber-type structure. While a role in tail formation and/or viral attachment is suggested, additional genetic and biochemical characterization of this group of proteins is needed.

While the large coiled-coil proteins are likely involved in the tail structure and assembly, the complex process of tail formation suggests that other proteins must be involved as well. ATSV shares a group of proteins with SMV1 and ATV (but not STSV1 and -2) that are potentially related to tail development. ATSV, SMV1, and ATV all contain a MoxR-like ATPase. In ATV, this ATPase has been implicated in tail formation along with three other virion proteins (two DNA binding proteins and the coiled-coil-containing protein gp48) (15). The functions of most MoxR ATPases are unknown, but some have been shown to be involved in various chaperone functions such as gas vesicle formation, cell envelope development, the formation of enzyme complexes, and stress response (15, 16). Additionally, ATSV, ATV, and SMV1 each have a protein containing a C-terminal metal-binding Von Willebrand factor A (VWA) domain (ATSV ORF22, ATV gp61, and SMV1 ORF759). This domain has been shown to interact with MoxR ATPases (17). Proteins containing the VWA domain are often found to mediate protein-protein interactions and enhance ATPase activity, including in ATV (15, 17). Even though ATSV and SMV1 have not been shown to produce tails extracellularly, it is possible that they contain a homologous chaperone-mediated mechanism of tail assembly using a related set of proteins, including the ones mentioned above. Since ATSV and ATV can develop longer tails (the average tail length of ATSV is 243 nm), perhaps these shared genes enable ATSV and SMV1 to develop longer tails than STSV1 and STSV2, in which tails average only 50 and 68 nm, respectively, in length.

The growing number of large spindle viruses and their interesting features led us to search for additional members in the environmental metagenomic data sets, using the MCP as a signature. Six viral and 20 cellular metagenomic data sets from Yellowstone hot springs (18–20) were queried by BLAST for the presence of large spindle virus-like virus MCP sequences. Relatives of ATSV (94 to 100% identity), ATV (65%), and SMV1 (70 to 82%) MCPs were found in DNA metagenomes from 5 hot springs (Fig. 1F). This analysis suggests that there are at least three groups of large spindle viruses present in Yellowstone's diverse hot spring environments. Given the diversity of large spindle viruses in Yellowstone, searching for large spindle viruses worldwide will no doubt also yield interesting new members and expand what we know about this fascinating group of viruses.

The growing number of isolated large spindle viruses has begun to reveal the defining features of this unique group of viruses. Their structural and genetic similarities necessitate their grouping into a superfamily. Further studies of these viruses and the addition of new members are sure to provide additional insight, not only for this new superfamily of archaeal viruses but for their respective hosts as well.

ACKNOWLEDGMENTS

This research was supported by National Science Foundation grants DEB-4W4596 (M.Y.), MCB-1413534 (C.M.L.), and MCB-0920312 (C.M.L.).

REFERENCES

1. **Lawrence CM, Menon S, Eilers BJ, Bothner B, Khayat R, Douglas T, Young MJ.** 2009. Structural and functional studies of archaeal viruses. J Biol Chem **284:**12599–12603. http://dx.doi.org/10.1074/jbc.R800078200.
2. **Dellas N, Snyder JC, Bolduc B, Young MJ.** 2014. Archaeal viruses: diversity, replication, and structure. Annu Rev Virol **1:**399–426. http://dx.doi.org/10.1146/annurev-virology-031413-085357.
3. **Prangishvili D, Vestergaard G, Haring M, Aramayo R, Basta T, Rachel R, Garrett RA.** 2006. Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. J Mol Biol **359:**1203–1216. http://dx.doi.org/10.1016/j.jmb.2006.04.027.
4. **Xiang XY, Chen LM, Huang XX, Luo YM, She QX, Huang L.** 2005. Sulfolobus tengchongensis spindle-shaped virus STSV1: virus-host interactions and genomic features. J Virol **79:**8677–8686. http://dx.doi.org/10.1128/JVI.79.14.8677-8686.2005.
5. **Erdmann S, Chen B, Huang XX, Deng L, Liu C, Shah SA, Bauer SL, Sobrino CL, Wang HN, Wei YL, She QX, Garrett RA, Huang L, Lin LB.** 2014. A novel single-tailed fusiform Sulfolobus virus STSV2 infecting model Sulfolobus species. Extremophiles **18:**51–60. http://dx.doi.org/10.1007/s00792-013-0591-z.
6. **Erdmann S, Bauer SL, Garrett RA.** 2014. Inter-viral conflicts that exploit host CRISPR immune systems of Sulfolobus. Mol Microbiol **91:**900–917. http://dx.doi.org/10.1111/mmi.12503.
7. **Haring M, Vestergaard G, Rachel R, Chen L, Garrett RA, Prangishvili D.** 2005. Virology: independent virus development outside a host. Nature **436:**1101–1102. http://dx.doi.org/10.1038/4361101a.
8. **Krupovic M, Quemin ERJ, Bamford DH, Forterre P, Prangishvili D.** 2014. Unification of the globally distributed spindle-shaped viruses of the Archaea. J Virol **88:**2354–2358. http://dx.doi.org/10.1128/JVI.02941-13.
9. **Goulet A, Vestergaard G, Felisberto-Rodrigues C, Campanacci V, Garrett RA, Cambillau C, Ortiz-Lombardia M.** 2010. Getting the best out of long-wavelength X-rays: de novo chlorine/sulfur SAD phasing of a structural protein from ATV. Acta Crystallogr D **66:**304–308. http://dx.doi.org/10.1107/S0907444909051798.
10. **Szymczyna BR, Taurog RE, Young MJ, Snyder JC, Johnson JE, Williamson JR.** 2009. Synergy of NMR, computation, and X-ray crystallography for structural biology. Structure **17:**499–507. http://dx.doi.org/10.1016/j.str.2009.03.001.
11. **Goulet A, Blangy S, Redder P, Prangishvili D, Felisberto-Rodrigues C, Forterre P, Campanacci V, Cambillau C.** 2009. Acidianus filamentous

virus 1 coat proteins display a helical fold spanning the filamentous archaeal viruses lineage. Proc Natl Acad Sci U S A **106:**21155–21160. http://dx.doi.org/10.1073/pnas.0909893106.
12. **Presnell SR, Cohen FE.** 1989. Topological distribution of 4-alpha-helix bundles. P Natl Acad Sci U S A **86:**6592–6596. http://dx.doi.org/10.1073/pnas.86.17.6592.
13. **Happonen LJ, Erdmann S, Garrett RA, Butcher SJ.** 2014. Adenosine triphosphatases of thermophilic archaeal double-stranded DNA viruses. Cell Biosci **4:**37. http://dx.doi.org/10.1186/2045-3701-4-37.
14. **Tao YZ, Strelkov SV, Mesyanzhinov VV, Rossmann MG.** 1997. Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. Structure **5:**789–798. http://dx.doi.org/10.1016/S0969-2126(97)00233-5.
15. **Scheele U, Erdmann S, Ungewickell EJ, Felisberto-Rodrigues C, Ortiz-Lombardia M, Garrett RA.** 2011. Chaperone role for proteins p618 and p892 in the extracellular tail development of Acidianus two-tailed virus. J Virol **85:**4812–4821. http://dx.doi.org/10.1128/JVI.00072-11.
16. **Iyer LM, Leipe DD, Koonin EV, Aravind L.** 2004. Evolutionary history and higher order classification of AAA plus ATPases. J Struct Biol **146:**11–31. http://dx.doi.org/10.1016/j.jsb.2003.10.010.
17. **Wong KS, Houry WA.** 2012. Novel structural and functional insights into the MoxR family of AAA+ ATPases. J Struct Biol **179:**211–221. http://dx.doi.org/10.1016/j.jsb.2012.03.010.
18. **Schoenfeld T, Patterson M, Richardson PM, Wommack KE, Young M, Mead D.** 2008. Assembly of viral metagenomes from Yellowstone hot springs. Appl Environ Microbiol **74:**4164–4174. http://dx.doi.org/10.1128/AEM.02598-07.
19. **Inskeep WP, Jay ZJ, Herrgard MJ, Kozubal MA, Rusch DB, Tringe SG, Macur RE, Jennings RD, Boyd ES, Spear JR, Roberto FF.** 2013. Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. Front Microbiol **4:**95. http://dx.doi.org/10.3389/fmicb.2013.00095.
20. **Bolduc B, Wirth J, Mazurie A, Young M.** 30 June 2015. Viral assemblage composition in Yellowstone acidic hot springs assessed by network analysis. ISME J http://dx.doi.org/10.1038/ismej.2015.28.